A PERTURBATION THEORY OF THE CONSTRAINED VARIATIONAL METHOD

IN MOLECULAR QUANTUM MECHANICS *

by

W. Byers Brown

University of Wisconsin Theoretical Chemistry Institute
Madison, Wisconsin

ABSTRACT

32 115

The variational problem of minimizing the energy of a trial wave function which is constrained to give the known theoretical or experimental expectation value of an operator is discussed. A perturbation approach is developed which leads to simple equations for estimating the effect of the constraint on the expectation values of other operators, and the increase in energy due to the constraint. It is shown that for hypervirial operators the constraint procedure for satisfying the corresponding hyper-virial theorem leads to equal and opposite results from the variational procedure. The theory is extended to cover the effect of constraints on second-order properties and multiple constraints. The perturbation equations are applied to Robinson's (1957) calculations on the lithium hydride molecule, previously discussed by Rasiel and Whitman (1964).

Author

- - - - -

INTRODUCTION

There has been a lot of interest recently in the calculation of the expectation values of operators other than the energy. It is well-known that if the error in a variationally determined wave function is of order Δ, then whereas the energy is correct to order Δ^2, the error in the expectation values of other operators is of order Δ *. An interesting approach introduced by Mukherji and Karplus[1] and developed recently by Whitman and his collaborators[2,3] is to constrain the variational wave function to give the known theoretical or experimental value μ of the expectation value M of some operator $\mathcal{M}$. This constraint will cost a certain amount of energy ΔE, but if the difference $\Delta\mu$ between μ and the free-variational value of M is small, then ΔE will only be of order $\Delta\mu^2$. The hope is that, for a negligible sacrifice in energy, the constrained wave function will lead to more accurate expectation values of other operators than the unconstrained wave function. This hope has been realized in calculations on hydrogen fluoride[1], lithium hydride[2] and helium[3].

The object of this paper is to present a new formulation of the constrained variational principle and to propose a perturbation approach based on the free variational principle. Rasiel and Whitman[2] took the first step towards such an approach but did not develop it.

- - - - -

* Unless the operator commutes with the Hamiltonian, in which case the error is also of order Δ^2.

1. A. Mukherji and M. J. Karplus, J. Chem. Phys., **38**, 44 (1963).
2. Y. Rasial and D. R. Whitman, Bull. Am. Phys. Soc., 9, 231 (1964); J. Chem. Phys. (to be published).
3. D. R. Whitman and R. Carpenter, Bull. Am. Phys. Soc., 9, 231 (1964).

CONSTRAINED VARIATIONAL PRINCIPLE

Let h be the Hamiltonian for the system of interest and let ψ and ϵ be the approximate ground state wave function and energy (non-degenerate) satisfying the familiar variational principle

$$\delta E = 0 \qquad \text{where} \qquad E = \langle \Psi, h\Psi \rangle / \langle \Psi, \Psi \rangle \qquad (1)$$

and Ψ is a trial wave function. The explicit form of the variational equation can be written

$$\langle \delta \Psi, (h - \epsilon) \psi \rangle = 0 . \qquad (2)$$

Consider now the constraint on Ψ : $M = \mu$ where

$$M = \langle \Psi, \mathcal{M} \Psi \rangle / \langle \Psi, \Psi \rangle \qquad (3)$$

and μ is a constant.[*] The constrained variational principle may be written

$$\delta E + \lambda \delta M = 0 \qquad (4)$$

where λ is a Lagrange multiplier, or in the more explicit form

$$\langle \delta \Psi, [h + \lambda(\mathcal{M} - \mu) - E] \Psi \rangle = 0 \qquad (5)$$

- - - - -

[*] It is assumed that μ is a possible value of M ; that is, does not exceed a bound of M. In the case of linear variation functions of the type (12) this is equivalent to assuming that μ lies between the smallest and largest eigenvalues of the matrix representative $\mathbf{M}$ of $\mathcal{M}$ in the basis set.

This equation is equivalent to the free variational principle $\delta \mathcal{E} = 0$ where

$$\mathcal{E} = \langle \Psi, \mathcal{H} \Psi \rangle / \langle \Psi, \Psi \rangle \tag{6}$$

and $\mathcal{H}$ is the fictitious Hamiltonian

$$\mathcal{H} = h + \lambda (\mathcal{M} - \mu) . \tag{7}$$

The value of the parameter λ is fixed by the constraint condition (3), which may be written

$$\langle \Psi, (\mathcal{M} - \mu) \Psi \rangle = 0 . \tag{8}$$

However, since λ does not occur explicitly in Ψ, the generalized Hellmann-Feynman theorem is valid for changes in λ[4]; that is

$$\frac{\partial \mathcal{E}}{\partial \lambda} = \langle \Psi, (\mathcal{M} - \mu) \Psi \rangle / \langle \Psi, \Psi \rangle . \tag{9}$$

Hence the constraint condition is equivalent to $\partial \mathcal{E} / \partial \lambda = 0$. The complete statement of the constrained variational principle is therefore that $\underline{\mathcal{E} \text{ is stationary with respect to variations in the}}$ $\underline{\text{trial wave function } \Psi \text{ and also with respect to variations in the}}$ $\underline{\text{Lagrange multiplier } \lambda}$. The stationary value of $\mathcal{E}$ is the

- - - - -
4. A. C. Hurley, Proc. Roy. Soc. <u>A226</u>, 179 (1954).

constrained energy E . For the ground state the Ψ extremum is

a minimum of $\mathcal{E}$ and the λ extremum is a maximum.

Let $E(\lambda)$ be the value of $\mathcal{E}$ which is stationary with

respect to $\delta\Psi$ for arbitrary λ . Then λ is determined by

$dE/d\lambda$ = 0 . The elimination of λ may be accomplished

formally by the following contour integral

$$E_{constrained} = \frac{1}{2\pi i} \oint \frac{E(\lambda) E''(\lambda)}{E'(\lambda)} d\lambda , \qquad (10)$$

where the primes denote differentiation and the contour goes round

the smallest zero of $E'(\lambda)$ in the complex λ-plane. Similarly,

if $L(\lambda)$ is the expectation value of an operator $\mathcal{L}$ for

arbitrary λ , then

$$L_{constrained} = \frac{1}{2\pi i} \oint \frac{L(\lambda) E''(\lambda)}{E'(\lambda)} d\lambda . \qquad (11)$$

Linear Variation Functions

If

$$\Psi = \sum_j c_j \phi_j \qquad (12)$$

where the ϕ_j are a fixed basis set with overlap matrix $\mathbf{S}$,

then equation (5) leads to the secular equations

$$(\mathbf{H} - E \mathbf{S}) \mathbf{c} = 0 \qquad (13)$$

where $\mathbb{H} = \mathbb{h} + \lambda (\mathbb{M} - \mu \mathbb{S})$ is the matrix representative of $\mathcal{H}$. Let

$$D(E, \lambda) \equiv \det \left\{ \mathbb{H} - E \mathbb{S} \right\} . \tag{14}$$

Then since

$$dE/d\lambda = - \frac{\partial D/\partial \lambda}{\partial D/\partial E} \quad (D = 0), \tag{15}$$

the first step in the constraint problem is to solve the equations

$$D(E , \lambda) = 0 \quad , \quad \partial D(E, \lambda)/\partial \lambda = 0 \quad (\partial D/\partial E \neq 0). \tag{16}$$

Rather than discuss further the nature of the solutions for this particular case, we turn to a general perturbation approach.

PERTURBATION APPROACH

Let us assume that $E (\lambda)$ can be expanded as a power series in λ over the range of interest, that is

$$E(\lambda) = \sum_{n=0}^{\infty} \lambda^n E^{(n)} . \tag{17}$$

The constraint condition is $dE/d\lambda = 0$ where

$$\frac{dE}{d\lambda} = M - \mu = \sum_{n=1}^{\infty} n \lambda^{n-1} E^{(n)} . \tag{18}$$

Let m be the expectation value of $\mathcal{M}$ for the unconstrained wave function, so that $M = m$ when $\lambda = 0$. If the error is $\Delta\mu = \mu - m$ then it follows from equation (18) by setting $\lambda = 0$ that

$$E^{(1)} = - \Delta\mu . \tag{19}$$

Hence the constraint condition for λ becomes

$$\Delta\mu = \sum_{n=2}^{\infty} n\lambda^{n-1} E^{(n)} . \tag{20}$$

Consider the case in which the unconstrained wave function ψ is a fairly good approximation to the true eigenfunction so that $\Delta\mu$ is small. Then by equation (20) λ is small and may be regarded as a perturbation parameter, in keeping with its role in the Hamiltonian $\mathcal{H}$ of equation (7). By inverting equation (20) λ may be expressed as a power series in $\Delta\mu$ which can be truncated when sufficient accuracy is achieved. Equation (17) can similarly be regarded as a perturbation series for the energy, and re-expressed in powers of $\Delta\mu$ if desired.

The first approximation for λ from equation (20) is

$$\lambda = \Delta\mu/2E^{(2)} + O(\Delta^2) . \tag{21}$$

The leading terms in the energy of constraint $\Delta E = E - \epsilon$ are

$$\Delta E = \lambda E^{(1)} + \lambda^2 E^{(2)} + O(\lambda^3) . \tag{22}$$

By substituting for $E^{(1)}$ from equation (19) and λ from equation (21) we get

$$\Delta E = - \frac{\Delta\mu^2}{2 E^{(2)}} + \frac{\Delta\mu^2}{4 E^{(2)}} + 0(\Delta^3) ,$$

or

$$\Delta E = - \Delta\mu^2/4 E^{(2)} + 0(\Delta^3) . \tag{23}$$

Note that since $E^{(1)}$ and λ are both of order $\Delta\mu$, while $E^{(2)}$ is of order unity in general, the terms formally of "first order" and "second order" are both of order $\Delta\mu^2$, and together are half the "first order" term $- \lambda \Delta\mu$.

Since $E^{(2)}$ is always negative for ground states it is clear from equation (23) that the constraint always leads to an increase in the calculated energy. However, this increase is only of order $\Delta\mu^2$ as expected, and will be small if $\Delta\mu$ is small and $\left| E^{(2)} \right|$ is large.

Consider the case when Ψ is the linear variational function (12), and let ψ_j, ϵ_j be the (j + 1)-th free variational solutions. In terms of this "unperturbed" basis the ground state (j = 0) expression for the second-order coefficient is

$$E^{(2)} = \sum_j{}' \frac{\left| M_{oj} \right|^2}{\epsilon_0 - \epsilon_j} . \tag{24}$$

It follows that $\left| E^{(2)} \right|$ is large if a low-lying excited state couples strongly with the ground state through the operator $\mathcal{M}$.

It can easily be shown that in equation (23) for ΔE the condition for the neglect of the term in $\Delta\mu^3$ compared to that in $\Delta\mu^2$ is

$$|\Delta\mu| \ll (E^{(2)})^2 / |E^{(3)}| . \qquad (25)$$

If $E^{(2)}$ and $E^{(3)}$ owe their magnitudes mainly to a low-lying excited state $j = 1$, then (25) reduces to

$$|\Delta\mu| = |\mu - M_{00}| \ll \frac{M_{01}^2}{|M_{11} - M_{00}|} . \qquad (26)$$

The terms in λ of higher order in $\Delta\mu$ are given in equation (74).

Other Expectation Values

Consider another operator $\mathcal{L}$ whose expectation value in the free variational approximation is $\ell = \langle \psi, \mathcal{L}\psi \rangle$, and let $L = \langle \Psi, \mathcal{L}\Psi \rangle$ be the constrained value. It is convenient to introduce another fictitious Hamiltonian

$$\tilde{H} = \tilde{h} + \gamma\mathcal{L} = h + \lambda(\mathcal{M} - \mu) + \gamma\mathcal{L}, \qquad (27)$$

and let E be the minimum variational energy, so that

$$L = (\partial E/\partial\gamma)_{\gamma=0} . \qquad (28)$$

Let us assume that E can be expanded as a double power series in λ and γ ,

$$E \;=\; \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} \lambda^{n} \gamma^{k} E^{(n,k)} \;. \tag{29}$$

Then

$$L \;=\; \sum_{n=0}^{\infty} \lambda^{n} E^{(n,1)} \;,\qquad \ell \;=\; E^{(0,1)} \;. \tag{30}$$

The change $\Delta L = L - \ell$ in the expectation value of $\mathcal{L}$ on constraining the variation is therefore given by

$$\Delta L \;=\; \sum_{n=1}^{\infty} \lambda^{n} E^{(n,1)} \;. \tag{31}$$

The first approximation for ΔL corresponding to equation (23) for ΔE is

$$\Delta L \;=\; \Delta\mu\, E^{(1,1)} / 2E^{(2)} \;+\; 0(\,\Delta^{2}) \;. \tag{32}$$

The leading change is therefore determined by the cross coefficient $E^{(1,1)}$ and will usually be of first order in $\Delta\mu$.

For the case of the linear variation function (12) the formula for $E^{(1,1)}$ is (assuming real functions),

$$E^{(1,1)} \;=\; 2 \sum_{j}{}' \frac{M_{0j}\,L_{j0}}{\epsilon_{0}-\epsilon_{j}} \;. \tag{33}$$

It is clear that $E^{(1,1)}$ will vanish, and the constraint will have little effect on the expectation value of $\mathcal{L}$, unless $\mathcal{M}$ and $\mathcal{L}$ interact in second order. They will not interact if the Hamiltonian

h possesses symmetry and the representation of the operator $\mathcal{ML}$ does not contain the unit representation.

Second Order Properties

The effect of constraints on second-order properties associated with an operator $\mathcal{L}$ can also be estimated by the perturbation approach. Second-order properties are those of the type[5]

$$Q = \frac{1}{2}\left(\partial^2 E / \partial \gamma^2 \right)_{\gamma=0}, \qquad (34)$$

such as electric polarizability. It follows from equation (29) that

$$Q = \sum_{n=0}^{\infty} \lambda^n E^{(n,2)}. \qquad (35)$$

The unconstrained value of Q corresponding to $\lambda = 0$ is therefore $q = E^{(0,2)}$, so that the change $\Delta Q = Q - q$ due to constraint is

$$\Delta Q = \sum_{n=1}^{\infty} \lambda^n E^{(n,2)}, \qquad (36)$$

$$= \Delta \mu E^{(1,2)}/2 E^{(2)} + 0(\Delta^2). \qquad (37)$$

For the linear variational function (12) the new coefficient appearing in this equation may be expressed in terms of the unconstrained basis set as

$$E^{(1,2)} = 2 \sum_{j}' \sum_{k}' \frac{L_{0j} \overline{M}_{jk} L_{k0} + 2 L_{0j} \overline{L}_{jk} M_{k0}}{(\epsilon_0 - \epsilon_j)(\epsilon_0 - \epsilon_k)} \qquad (38)$$

5. J. O. Hirschfelder, W. Byers Brown and S. T. Epstein, _Advances in Quantum Chemistry_, volume 1, ed. Löwdin, Academic Press, New York(1965).

where $\overline{M}_{jk} = M_{jk} - M_{oo} \delta_{jk}$, $\overline{L}_{jk} = L_{jk} - L_{oo} \delta_{jk}$.

It can be seen from the equations of this section how easy it would be to generalize the theory to cover the <u>constraint</u> of a second-order property. However , such a procedure is unlikely to be useful in practice, and will not be developed here.

CONSTRAINT COEFFICIENTS

The coefficients $E^{(2)}$, $E^{(1,1)}$ and $E^{(1,2)}$ appearing in the leading terms for the changes in E, L and Q due to the constraint $M = \mu$ refer to the variational approximation under consideration. However, if the approximation is a good one, which is a necessary condition for the success of the perturbation treatment, then the values of the coefficients given by the variational approximation will be close to the exact values.

To exhibit the relation of the exact and approximate coefficients, suppose P is the projection operator onto the subspace of the variational function Ψ , so that in terms of the unconstrained orthonormal basis set ψ_j

$$P = \sum_j |\psi_j\rangle\langle\psi_j| . \tag{39}$$

Also let O be the part of P complementary to $\psi \equiv \psi_o$, the variational approximation to the ground state eigenfunction of h, so that

$$O = P - |\psi\rangle\langle\psi| = \sum_j{}' |\psi_j\rangle\langle\psi_j| . \tag{40}$$

The approximate coefficients may then be written in the form

$$E^{(2)} = \langle \psi, \mathcal{M} \mathcal{G} \mathcal{M} \psi \rangle \quad , \tag{41}$$

$$E^{(1,1)} = 2 \langle \psi, \mathcal{L} \mathcal{G} \mathcal{M} \psi \rangle \quad , \tag{42}$$

$$E^{(1,2)} = 2 \Big[\langle \psi, \mathcal{L} \mathcal{G} \mathcal{L} \mathcal{G} \mathcal{M} \psi \rangle +$$

$$\langle \psi, \mathcal{L} \mathcal{G} \mathcal{M} \mathcal{G} \mathcal{L} \psi \rangle + \langle \psi, \mathcal{M} \mathcal{G} \mathcal{L} \mathcal{G} \mathcal{L} \psi \rangle \Big], \tag{43}$$

where

$$\mathcal{G} = \mathcal{O} (\epsilon - \tilde{h})^{-1} \mathcal{O} \quad , \quad \tilde{h} = P h P . \tag{44}$$

The exact coefficients are given formally by the same expressions but now $P \to 1$, so $\tilde{h} \to h$ and $\mathcal{O} \to 1 - |\psi\rangle\langle\psi|$, and ψ, ϵ are now the ground state eigenfunction and eigenvalue of h.

For certain operators $\mathcal{M}$ and $\mathcal{L}$ the exact coefficients involve only familiar physical properties. If these are known experimentally or theoretically it should be easy to make estimates of the effects of the corresponding constraints on certain of the calculated properties. This is illustrated by the following examples for molecular systems.[*]

- - - - -

[*] In this section the Born-Oppenheimer fixed nucleus approximation is assumed, and all formulae are in conventional atomic units.

a) <u>Dipole Moment</u>

If $\mathcal{M}$ is the dipole moment vector operator for a polyatomic molecule,

$$\mathcal{M} = \sum_\alpha Z_\alpha \underset{\sim}{r}_\alpha - \sum_i \underset{\sim}{r}_i \; , \tag{45}$$

the exact coefficient $E^{(2)} = -\frac{1}{2} \underset{\approx}{\alpha}$, where $\underset{\approx}{\alpha}$ is the electric polarizability tensor. The increase in energy on constraining the dipole moment to its correct value $\underset{\sim}{\mu}$ is therefore approximately

$$\Delta E \cong \frac{1}{2} \Delta\underset{\sim}{\mu} \cdot \underset{\approx}{\alpha}^{-1} \cdot \Delta\underset{\sim}{\mu} \; ,$$

$$= \Delta\mu^2 / 2\alpha_{\mu\mu} \tag{46}$$

where $\alpha_{\mu\mu}$ is the diagonal component of $\underset{\approx}{\alpha}$ along the direction of $\Delta\underset{\sim}{\mu}$, assumed to be a principal axis.

b) <u>Nuclear Force</u>   (Hellmann-Feynman Theorem)

If $\mathcal{M}$ is the force vector operator for nucleus α in a molecule

$$\underset{\sim}{\mathcal{F}}^\alpha = - Z_\alpha \sum_{\beta \neq \alpha} \frac{Z_\beta \underset{\sim}{R}_{\alpha\beta}}{R_{\alpha\beta}^3} + Z_\alpha \sum_i \frac{\underset{\sim}{r}_{\alpha i}}{r_{\alpha i}^3} \; , \tag{47}$$

then by the Hellmann-Feynman theorem the expectation value of $\underset{\sim}{\mathcal{F}}^\alpha$ should vanish for all nuclei in the equilibrium configuration. The exact second-order tensor coefficient corresponding to this operator appears in the formula for the force-constant matrix of the molecule.[6]

- - - - -

6. W. Byers Brown, Proc. Camb. Phil. Soc., <u>54</u>, 251 (1958).

For a diatomic molecule AB the zz-component along the molecular axis can be written in the form[7,8]

$$E_\alpha^{(2)} = \frac{1}{2} k - \frac{1}{2} z_\alpha \left(\frac{4\pi}{3} \rho_\alpha + q_\alpha \right) \quad (\alpha = A, B) \quad (48)$$

where ρ_α is the electron density at nucleus α, k is the force constant and q_α is the zz-component of the field gradient at nucleus α. Unfortunately, although k and q_α may be accessible experimentally, ρ_α usually is not.

The expectation value of the total force operator

$$\underset{\sim}{7} = \sum_\alpha \underset{\sim}{7}^\alpha \quad (49)$$

should be zero for <u>all</u> configurations. The corresponding second-order coefficient for a diatomic molecule has the zz-component

$$E^{(2)} = \frac{2 z_A z_B}{R^3} - \frac{1}{2} \sum_\alpha z_\alpha \left(\frac{4\pi}{3} \rho_\alpha + q_\alpha \right) \quad (50)$$

where R is the internuclear distance AB.

The second order cross coefficient $E^{(1,1)}$ for $\mathcal{M}$ = dipole moment and $\mathcal{L}$ = total nuclear force 7_z is particularly simple. By the oscillator strength sum rule the zz-component is simply

$$E^{(1,1)} = N \quad (51)$$

- - - - -

7. Erich Steiner, Ph.D. Thesis, University of Manchester, England (1961).
8. Lionel Salem, J. Chem. Phys., <u>38</u>, 1227 (1963).

where N is the number of electrons. If $\mathcal{L}$ is instead only the force $\mathcal{F}_z^\alpha$ on nucleus α then

$$E_\alpha^{(1,1)} = Z_\alpha \mp \partial\mu/\partial R \qquad (52)$$

where μ is here the dipole moment. In equation (52) the positive sign is to be taken for the nucleus with the smaller z-coordinate and the negative sign for the other nucleus. It has been assumed that the molecule is electrically neutral so that $N = Z_A + Z_B$.

c) Virial Theorem

According to the virial theorem the expectation value of the operator $\mathcal{H} = 2\mathcal{K} + \mathcal{U}$, where $\mathcal{K}$ and $\mathcal{U}$ are the kinetic and potential energy operators, should be equal to $\mu = -\sum_\alpha R_\alpha \cdot \partial\epsilon/\partial R_\alpha$ *. If this condition is not satisfied the wave function may be constrained to do so if desired. The cost in energy can be estimated from equation (23) by using the exact expression for the second-order coefficient which is[6,7]

$$E^{(2)} = \epsilon + 2\sum_s R_s(\partial\epsilon/\partial R_s) + \tfrac{1}{2}\sum_s\sum_t R_s k_{st} R_t \qquad (53)$$

where R_s is the length of bond s and k_{st} are the bond stretching force constants.

- - - - -

* This is not strictly the original virial theorem, but a closely related theorem derived by homogeneous scaling. The original virial theorem leads to

$$2K + U = -\sum_\alpha R_\alpha \cdot \langle \psi, \partial\mathcal{H}/\partial R_\alpha \psi \rangle$$

and is only equivalent to the scaled form quoted in the text if the Hellmann-Feynman theorem is valid.

The conclusion that the energy must <u>increase</u> in order to satisfy the virial theorem appears wrong at first sight. It is well known that an approximate wave function which does not satisfy the virial theorem can always be made to do so by introducing a scale parameter and minimizing the expectation value of the Hamiltonian with respect to it. This procedure always leads to a <u>decrease</u> in energy. The paradox is resolved by observing that if the variational function included coordinate scaling among its degrees of variational freedom, then the unconstrained function ψ would already satisfy the virial theorem, and no further constraint would be necessary. If the virial theorem is not satisfied by ψ this implied that the scale in Ψ is not regarded as a possible variation.

The intimate relationship between the effects of <u>constraining</u> an approximate wave function to satisfy a theoretical condition and the effects of <u>varying</u> the wave function to satisfy the condition are discussed in the next section for the case in which the condition can be written as a hypervirial relation.

d) <u>Molecular Geometry</u>

It is also possible to estimate the effect of a constraint condition on the equilibrium configuration calculated for a molecule. Consider the simplest case of a diatomic molecule and let r_e be the equilibrium bond length for the unconstrained treatment. Impose the constraint

$$M(R) = \mu(R) \tag{54}$$

for every internuclear distance R. Let R_e be the experimental

bond length satisfying the equation $\partial E / \partial R = 0$, or

$$\frac{\partial \Delta E}{\partial R} + \frac{\partial \epsilon}{\partial R} = 0 \qquad (R = R_e) \tag{55}$$

where ΔE is the energy of constraint given by equations (22) and

(23). Since $\Delta R = R_e - r_e$ is supposed small we can expand to

get

$$\left(\frac{\partial \epsilon}{\partial R}\right)_{R=R_e} = k_e \Delta R + O(\Delta R^2), \tag{56}$$

where k_e is the unconstrained force constant. Therefore,

$$\Delta R = -\frac{1}{k_e} \frac{\partial \Delta E}{\partial R} + \ldots, \tag{57}$$

and since ΔE is of order $\Delta \mu^2$, the error in R_e is

effectively of the same order.

CONSTRAINING VERSUS VARYING FOR HYPERVIRIAL THEOREMS

The virial theorem and the Hellman-Feynman theorem for the

total force are particular cases of a general class known as the

hypervirial theorems.[9,10] These are relations of the form

$$\langle \psi, [h, W] \psi \rangle = 0, \tag{58}$$

- - - - -

9. J. O. Hirschfelder, J. Chem. Phys. <u>33</u>, 1762 (1960).
10. S. T. Epstein and J. O. Hirschfelder, Phys. Rev. <u>123</u>, 1495 (1960).

and are satisfied by the exact wave function for any operator $\mathcal{W}$;
in what follows we shall suppose that $\mathcal{W}$ is hermitian. The
hypervirial theorem (58) can be regarded as the theoretical condition
$\mu = 0$ on the expectation value of the hermitian operator $\mathcal{M}$
defined by

$$\mathcal{M} = i\left[\mathcal{h}, \mathcal{W}\right] . \tag{59}$$

Epstein and Hirschfelder[10] have shown that an approximate wave
function can always be made to satisfy any particular hypervirial
theorem by a variational procedure. This is done by introducing a
certain mode of variation depending on $\mathcal{W}$ and making the energy
stationary with respect to this mode. If we wish to modify an
approximate wave function to satisfy a particular hypervirial
theorem we have, therefore, a choice: (a) constrain the wave
function to satisfy $\mu = 0$; (b) vary the wave function to
satisfy $\mu = 0$. The theory of the constraint procedure (a) and
its perturbation development have been discussed in the preceding
section. In this section a corresponding discussion of the
variational procedure (b) will be given. It turns out that the
procedures are intimately related in that the leading terms in the
perturbation treatments are equal and opposite. In order to
exhibit the relationship clearly, the constraint procedure will be
summarized first, using a slightly modified notation.

a) Constraint Method

To minimize $E = \langle \Psi, h\Psi \rangle / \langle \Psi, \Psi \rangle$ subject to

$$\mu = \langle \Psi, \mathcal{M}\Psi \rangle / \langle \Psi, \Psi \rangle = 0 \tag{60}$$

we define

$$\mathcal{H}_\lambda = h + \lambda \mathcal{M} \tag{61}$$

and let

$$E_\lambda = \langle \Psi, \mathcal{H}_\lambda \Psi \rangle / \langle \Psi, \Psi \rangle \tag{62}$$

be the minimum value of the expectation value of $\mathcal{H}_\lambda$ with respect to variations $\delta\Psi$. The value of the Lagrange multiplier λ for which (60) is valid is given by

$$dE_\lambda / d\lambda = 0 . \tag{63}$$

Expanding E_λ in a perturbation series gives

$$E_\lambda = \epsilon + \lambda m + \lambda^2 E^{(2)} + \cdots \tag{64}$$

where $m = \langle \psi, \mathcal{M}\psi \rangle = E_\lambda^{(1)}$ is the expectation value of $\mathcal{M}$ for the unconstrained wave function. Assuming m is small, it follows from equation (63) that

$$\lambda = -m/2E_\lambda^{(2)} + \cdots , \tag{65}$$

and therefore the energy loss due to constraining is

$$\Delta E_\lambda = E_\lambda - \epsilon = -m^2/4 \, E_\lambda^{(2)} + \ldots \quad . \tag{66}$$

The second-order coefficient $E_\lambda^{(2)}$ is given by equation (41). Substituting for $\mathcal{M}$ from equation (59), and replacing the approximate coefficient by the exact one, equation (41) becomes

$$E_\lambda^{(2)} = - \langle W\psi, (h - \epsilon) W\psi \rangle \leq 0 . \tag{67}$$

Since $E_\lambda^{(2)}$ is clearly negative, this confirms that ΔE_λ is positive.

The change in an expectation value ΔL_λ due to constraining is given by equation (32), or

$$\Delta L_\lambda = -m E_\lambda^{(1,1)}/2 \, E_\lambda^{(2)} + \ldots \quad . \tag{70}$$

The exact form of $E_\lambda^{(1,1)}$ when $\mathcal{M}$ has the commutator form (59) is

$$E_\lambda^{(1,1)} = \langle \psi, i[\mathcal{L}, W] \psi \rangle . \tag{71}$$

b) Variation Method

The hypervirial theorem for W is satisfied by the variational function

$$\Psi_\eta = e^{i\eta W} \psi \tag{72}$$

if the trial energy

$$E_\eta = \langle \Psi_\eta, \hbar \Psi_\eta \rangle / \langle \Psi_\eta, \Psi_\eta \rangle \tag{73}$$

is stationary with respect to the parameter η [10]. Substituting equation (72) into equation (73) we get

$$E_\eta = \langle \psi, \hbar_\eta \psi \rangle \tag{74}$$

where

$$\hbar_\eta = e^{-i\eta w} \hbar \, e^{i\eta w},$$

$$= \hbar + \eta \, i[\hbar, w] + \tfrac{1}{2}\eta^2 [w,[\hbar,w]] + \cdots \tag{75}$$

The perturbation expansion of E_η is therefore

$$E_\eta = \epsilon + \eta m + \eta^2 E_\eta^{(2)} + \cdots \tag{76}$$

where

$$E_\eta^{(2)} = \tfrac{1}{2} \langle \psi, [w,[\hbar,w]] \psi \rangle. \tag{77}$$

If $E_\eta^{(2)}$ is replaced by the coefficient in which ψ is an exact eigenfunction of $\hbar$ with eigenvalue ϵ, then

$$E_\eta^{(2)} = \langle w\psi, (\hbar - \epsilon) w\psi \rangle \geqslant 0,$$

$$= - E_\lambda^{(2)}. \tag{78}$$

Setting $dE_\eta / d\eta = 0$ and assuming m is small we get

$$\eta = -m/2E^{(2)} + \ldots , \qquad (79)$$

and therefore the energy <u>gain</u> due to varying is

$$\Delta E_\eta = -m^2/ 4 E_\eta^{(2)} + \ldots , $$

$$= - \Delta E_\lambda + \ldots \qquad (80)$$

The energy gain on varying is thus equal and opposite to the energy loss on constraining, to the leading order in perturbation theory.

The expectation value of an operator $\mathcal{L}$ is given by

$$L_\eta = \langle \Psi_\eta, \mathcal{L} \Psi_\eta \rangle / \langle \Psi_\eta, \Psi_\eta \rangle ,$$

$$= \langle \psi, \mathcal{L}_\eta \psi \rangle \qquad (81)$$

where

$$\mathcal{L}_\eta = e^{-i\eta W} \mathcal{L} e^{i\eta W},$$

$$= \mathcal{L} + \eta i [\mathcal{L}, W] + O(\eta^2). \qquad (82)$$

Hence the change due to the variation is

$$\Delta L_\eta = \eta \langle \psi, i[\mathcal{L}, W] \psi \rangle + O(\eta^2),$$

$$= - m \frac{\langle \psi, i[\mathcal{L}, W] \psi \rangle}{2 E_\eta^{(2)}} + \ldots,$$

$$= - \Delta L_\lambda + \ldots \qquad (83)$$

The leading changes in expectation values are therefore also equal and opposite for the two methods. This result suggests that it would be desirable to choose an approximate ψ so that the coefficient $E^{(1,1)}$ vanishes; that is, that the expectation value of the commutator of $\mathcal{L}$ and W vanishes. However, this is not in general true for the exact wave function. If it were true, then expectation values would possess a stationary property, which alas is not so.

We are therefore left with the somewhat unsatisfactory situation that "yer pays yer money and yer takes yer choice". It can, of course, be argued that the η-variation procedure outlined above is artificial. The variational freedom implicit in the approximate wave function ψ has been frozen while varying η. If these variations were allowed to interact, the changes ΔL in expectation values might turn out to have the same sign. However, in this case detailed information is needed about the effect of the hypervirial operator W on the subspace of the variational function described by P, and it is not possible to estimate changes in the simple way developed above.

MULTIPLE CONSTRAINTS

The theory can be generalized to cover the imposition of many constraints of the form

$$M_\xi = \langle \Psi, \mathcal{M}_\xi \Psi \rangle / \langle \Psi, \Psi \rangle = \mu_\xi \quad (\xi = 1, 2, .., \nu). \quad (84)$$

Let $\mathcal{E}$ be the expectation value of the fictitious Hamiltonian

$$\mathcal{H} = h + \sum_{\zeta} \lambda_{\zeta} (\mathcal{M}_{\zeta} - \mu_{\zeta}) \qquad (85)$$

for a variational wave function Ψ. Then the generalized constrained variational principle is that $\mathcal{E}$ is stationary with respect to variations in Ψ (minimum) and with respect to variations in all the λ_{ζ} (maximum). Alternatively, if $E(\lambda_1, \lambda_2, \ldots, \lambda_{\nu})$ is the value of $\mathcal{E}$ which is stationary with respect to $\delta \Psi$ for arbitrary λ_{ζ}, then the constraint conditions determining the latter are

$$\partial E / \partial \lambda_{\zeta} = 0 \qquad (\zeta = 1, 2, \ldots, \nu). \qquad (86)$$

The perturbation series for the energy change due to the constraints is

$$\Delta E = \sum_{\zeta} \lambda_{\zeta} E_{\zeta} + \sum_{\zeta} \sum_{\eta} \lambda_{\zeta} \lambda_{\eta} E_{\zeta\eta} + O(\lambda^3) \qquad (87)$$

where the superscripts denoting the order have been omitted. Since $\partial \Delta E / \partial \lambda_{\zeta} = M_{\zeta} - \mu_{\zeta}$ it follows from equation (87) that

$$E_{\zeta} = - \Delta \mu_{\zeta} \qquad (88)$$

where $\Delta \mu_{\zeta} = \mu_{\zeta} - m_{\zeta}$ is the error in the free variational expectation value of $\mathcal{M}_{\zeta}$. The Lagrange multipliers are therefore given by the equations

$$\Delta \mu_{\zeta} = 2 \sum_{\eta} \lambda_{\eta} E_{\zeta \eta} + O(\lambda^2) \quad (\zeta = 1, 2, .., \nu). \quad (89)$$

Hence, solving approximately for the λ_{ζ} and substituting into equation (87) gives

$$\Delta E = -\frac{1}{4} \sum_{\zeta} \sum_{\eta} \Delta \mu_{\zeta} E_{\zeta \eta}^{-1} \Delta \mu_{\eta} + O(\Delta^3), \quad (90)$$

where $E_{\zeta \eta}^{-1}$ is the $\zeta \eta$-element of the inverse of the matrix of elements $E_{\zeta \eta}$. The change ΔL in an unconstrained expectation value L due to multiple constraints is similarly easily shown to be

$$\Delta L = \sum_{\zeta} \sum_{\eta} \Delta \mu_{\zeta} E_{\zeta \eta}^{-1} E_{\eta L} + O(\Delta^2). \quad (91)$$

For the linear variational function (12) the coefficients occurring in equations (90) and (91) are given by

$$E_{\zeta \eta} = \sum_{j}' \frac{(M_{\zeta})_{oj} (M_{\eta})_{jo}}{\epsilon_o - \epsilon_j} \quad, \quad E_{\eta L} = \sum_{j}' \frac{(M_{\eta})_{oj} L_{jo}}{\epsilon_o - \epsilon_j} \quad. \quad (92)$$

ALTERNATIVE PERTURBATION SCHEMES

The perturbation approach developed above corresponds to the Rayleigh-Schrödinger perturbation theory. It is the simplest perturbation scheme and has the merit of providing explicit formulae for quantities of interest. However, in practice it may happen that although the conditions for a perturbation approach

are satisfied, in that ΔE and $\Delta \mu$ are small and $\Delta E = 0(\Delta \mu^2)$, the Rayleigh-Schrödinger series converges too slowly to be useful. This can occur if there is a low-lying excited state for which the expectation value of $\mathcal{M}$ differs considerably from the ground state value, M_{oo}. In such cases the alternative perturbation schemes of Brillouin-Wigner[11], Feinberg[11] or Sasakawa[12] or the partitioning technique of Löwdin[13] may converge more rapidly, and may be used when the variational function is of the type (12).

The structure of the Brillouin-Wigner series is particularly simple; the first terms in ΔE are

$$\Delta E = - \lambda \Delta \mu - \lambda^2 \sum_j \frac{|M_{oj}|^2}{\epsilon_j - E + \lambda(M_{jj} - \mu)} + O(\lambda^3). \quad (93)$$

In principle this must be solved iteratively. However, ΔE is of order λ^2, so for simplicitly the approximation (not exact to order λ^2)

$$E \cong \epsilon_o - \lambda \Delta \mu \quad (94)$$

may be used in the denominators in equation (93) to give

$$\Delta E = - \lambda \Delta \mu - \lambda^2 \sum_j{}' \frac{|M_{oj}|^2}{\epsilon_{jo} + \lambda \bar{M}_{jj}} + O(\lambda^3) \quad (95)$$

- - - - -

11. See A. Dalgarno, "Stationary Perturbation Theory", Chap. 5, 171 and D. R. Bates (ed.) Quantum Theory, 1, Academic Press, New York (196).

12. Sasakawa, J. Math. Phys. 4, 970 (1963).

13. P. O. Löwdin, J. Mol. Spec. 10, 12 (1963); ibid., 14, 112 (1964).

where $\epsilon_{jo} = \epsilon_j - \epsilon_o$. The Lagrange parameter λ can be obtained from the equation $d\Delta E / d\lambda = 0$ by solving iteratively, in the spirit of the Brillouin-Wigner theory.

APPLICATION TO LITHIUM HYDRIDE

To illustrate the application of the perturbation approach and also to examine its range of validity, the theory has been applied to Robinson's[14] calculations on lithium hydride, which were chosen by Rasiel and Whitman[2] in their study of the constrained variational method. Robinson's work is based on a three-term open-shell configuration interaction function with a basis set of six Slater type atomic orbitals. The minimum total molecular energy obtained by Robinson differs from the experimental value of -219.71 eV by only 0.47 eV, and the equilibrium distance is 3.0133 Bohr, which agrees with the experimental value.[*] However, the expectation values of operators are not nearly so good. In particular, the calculated dipole moment is 4.157 D (H $\longrightarrow$ Li) which differs from the experimental value of 5.881 D by 29 per cent; the expectation value of the total force on the molecule, which should vanish by the Hellmann-Feynman theorem, is 0.698 au. (H $\longrightarrow$ Li); the quantity $2K + U$, which should vanish by the virial theorem, is -0.550 au.

- - - - -

14. J. M. Robinson, Jr., Ph.D. Thesis, University of Texas, Austin (1957).

* Robinson's work has been surpassed in accuracy by the recent 28-term calculation, using a mixed orbital-basis set, of J. C. Browne and F. A. Matsen, Phys. Rev. <u>135</u>, A1227 (1964).

Rasiel and Whitman[2] constrained Robinson's variational function to reproduce the experimental dipole moment by setting $\Delta\mu = 1.724$ D $= 0.678$ au. They calculated an increase in energy ΔE due to the constraint of only 0.00506 au. or about 0.14 eV. The expectation values of the operators they considered improved quite markedly in the cases where the value is known. In particular the constrained expectation value of the total force was reduced to 0.277 au.

In order to check the range of validity of the perturbation approach, the calculations performed by Rasiel and Whitman were repeated using the formulae derived in this paper. The value -29.66 au. for the coefficient $E^{(2)}$ was obtained from equation (24) by substituting the matrix elements M_{oj} calculated from Robinson's results by Rasiel.[15,*] The first approximation for ΔE, equation (23), then gives 0.00387 au. or about 0.11 eV compared with the correct value 0.14 eV found by Rasiel and Whitman. The inequality (26) can be applied in this case to find out if the higher terms can be neglected, as LiH possesses a low-lying excited state. The inequality becomes: $0.68 \ll 1.55$. Since this is not strictly

- - - - -

15. Y. Rasiel, Ph.D. Thesis, Case Institute of Technology (1964), and private communication.

* This corresponds to a value of the polarizability parallel to the LiH axis of 8.8 A^3, which is over twice as large as the approximate value 3.76 A^3 calculated by H. J. Kolker and M. Karplus, J. Chem. Phys. $\underline{39}$, 2011 (1963). These authors think their method already overestimates α, so that the replacement of $E^{(2)}$ by its experimental value would be in error by over a factor of two.

satisfied, the convergence of the perturbation series was checked
as follows. λ was calculated by inverting equation (20) correctly
through third order in $\theta = \Delta\mu/2E^{(2)}$:

$$\lambda = \theta - (3E^{(3)}/2E^{(2)})\theta^2 + 2\left[(3E^{(3)}/2E^{(2)})^2 - (E^{(4)}/E^{(2)})\right]\theta^3 + O(\theta^4)$$

(96)

$$= -0.0114 - 0.0034 - 0.0002 + \ldots ,$$

$$= -0.0150 .$$

This value was then substituted into equation (17) through terms
in λ^4 :

$$\Delta E = -\lambda\Delta\mu + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \lambda^4 E^{(4)} + O(\lambda^5) , \quad (97)$$

$$= +0.01017 - 0.00667 + 0.00175 - 0.00020 + \ldots ,$$

$$= +0.00504 .$$

The final agreement with Rasiel and Whitman is excellent, but the
convergence of the Rayleigh-Schrödinger series is only moderately
fast. The first order correction to the total force F due to
constraint was calculated from equation (32). The appropriate
$E^{(1,1)} = 49.60$ au.* was obtained by substituting the matrix

- - - - -

* The exact value of the coefficient $E^{(1,1)}$, given by equation (51),
is 4   for LiH. This is a warning that the use of theoretically
or experimentally known values of the exact coefficients in
estimating the effects of constraints may be very wide of the mark.

elements for M_{oj} and F_{oj} calculated by Rasiel[13] into equation (33). This leads to $\Delta F \cong -0.567$ or F(constrained) $\cong 0.131$ au., which is to be compared with Rasiel and Whitman[2]: F(constrained) = 0.277 au. In this case the leading term in the perturbation formula gives the correct sign and order of magnitude for ΔL, but higher order terms would be needed to compute it with precision.

The particular constraint used by Mukherji and Karplus[1] and by Rasiel and Whitman[2], namely, that the dipole moment shall have the observed value, spoils the absolute character of the theoretical calculation by introducing an empirical element. This criticism cannot be made of the constraint that the total force F vanishes, since it is a theoretical requirement of the Hellmann-Feynman theorem. It is therefore interesting to use the perturbation formulae to find the cost in energy ΔE of the constraint F = 0, and its effect on the calculated dipole moment. The quantity $\Delta \mu$ is now $\Delta F = -0.698$ av. and the appropriate $E^{(2)} = -65.96$, so that substitution in equation (23) gives $\Delta E = 0.00184$ au. or about 0.05 eV. Pursuing the terms of higher order to check convergence, the value of λ given by equation (96) is

$$\lambda = +0.00529 - 0.00089 - 0.00054 + \ldots ,$$

$$= +0.00386 .$$

When this value is substituted into equation (97) along with the appropriate values for the coefficients $E^{(2)}$, etc., the result is

of the same symmetry or similar functional behaviour are applied any

approximate wave function will inevitably become distorted and

unreliable. However, future interest is focussed on the possibility

of applying only a few different types of theoretical constraint

and calculating the effect on a wide variety of other properties.

ACKNOWLEDGEMENTS

$$\Delta E = +0.00269 - 0.00098 - 0.00008 - 0.00004 + \dots,$$

$$= +0.00159 \text{ au.}$$

or about 0.04 eV. The convergence is therefore satisfactory, and the first approximation for the change ΔM in the dipole moment given by equation (32) may be presumed dominant. The value of $E^{(1,1)}$ is the same as before, and the result is $\Delta M \cong 0.262$ au. or M (constrained) $\cong 4.83$ D, which is in error by about 18 per cent. The effect of constraining the total force is, therefore, to improve the calculated dipole moment by about 9 per cent[*], for a negligible change in energy of about 0.04 eV. The improvement is not striking, but neither is it negligible.

DISCUSSION

The perturbation approach presented here should provide a rapid and easy way to estimate possible improvements to be gained by modifying variational wave functions and the cost thereof. However, the application to lithium hydride shows, first, that attention must be paid to the convergence of the series to guarantee reliability; and second, that the substitution of the exact values of the perturbation coefficients for the variational values may be unreliable.

With regard to multiple constraints, clearly as more constraints

- - - - -

[*] Y. Rasiel (private communication) has calculated ΔE and ΔM by the iterative method of reference 2. He obtains $\Delta E = + 0.00150$ au. or 0.04 eV, and M(constrained) = 4.75 D. The agreement is thus very good.